

Eur. J. Clin. Chem. Clin. Biochem.
Vol. 31, 1993, pp. 851–859

© 1993 Walter de Gruyter & Co.
Berlin · New York

A European Quality Control Programme as a Cooperative Tool between Users and a Diagnostic Company

By G. M. P. J. Costongs and P. C. W. Janson

Department Clinical Chemistry, Maaslandziekenhuis, Sittard, The Netherlands

(Received February 24/September 28, 1993)

Summary: Achievement of appropriate precision is a major challenge for the clinical laboratory, particularly in specialized fields such as enzymology and immunochemistry. Especially in immunochemistry, where accuracy can only poorly be checked because of a lack of reference methods and a multitude of detectable epitopes, precision remains an important index of quality. In immunochemical methods there are three main sources of variation: differences in methodology, lot-to-lot variation and performance of the assay, the latter including the skill of the technicians and instrument performance. The results of a European Quality Control Programme for the users of one type of automated immunoassay analyser were compared with a similar Italian Quality Control Programme for users of different immunometric techniques. The precisions (%CV) of the two programmes were compared, as well as the number of values lying within one half of the biological variation of the analysed serum components. Feedback between users and manufacturers leads to a clear improvement in analysis performance, so that most of the considered values come to lie within the acceptable limits of $\frac{1}{2}$ of the biological variation. It is concluded that both national and international quality control programmes are very useful for indicating the intrinsic quality of the quantities used in the clinical laboratory.

Introduction

The achievement of appropriate precision is a major challenge for the clinical laboratory today. This applies to general clinical chemistry, and to more specialized fields, such as enzymology and immunochemistry. In immunochemistry accuracy can only poorly be checked because of a lack of internationally accepted reference methods and because a multitude of epitopes is detected by different commercially available tests. Therefore precision remains an important index of quality in this field.

At the moment, the precision of immunoassays is still less than that of other clinical chemical methods. Three main sources can be found for this variation:

1. the difference in technique, arising from the use of distinctive antibodies and adjuvant reagents;
2. the imprecision of the reagents, which is related to lot-to-lot variations, and between/within run os-

cillations. This is mainly a problem of radioimmunoassays, which must use relatively small batches due to their limited shelf-life;

3. the performance of the method/reagents used, which are a result of the intrinsic performance of the technology, the skill of the laboratory technicians and the precision of the instruments used.

The above considerations imply that when the precision of an immunoassay is to be enhanced, this can best be done when manufacturers collaborate with the users of their reagents and or instruments. With regard to the different causes of imprecision, the following can be said.

1. Although differences in techniques are not the immediate responsibility of the manufacturer, international organisations can play an important coordinating role.

2. Lot-to-lot variations can be reduced by producing larger batches and by using non-isotopic reagents with longer shelf-lives.
3. Performances can be improved by efficient feedback between user and manufacturer.

On the other hand, the concentration of serum constituents in human beings is subject to individual variations (1–3). In the seventies, when only isotopic tests with high variations were available, authors started to consider the combination of biological and analytical variation (1–6). Thus, one of the goals of an immunoassay should be to reduce the variation of a measured value to $\frac{1}{2}$ of the biological variation ($\frac{1}{2}$ BV), and to reduce it sufficiently for the needs of medical decision making (7).

At the end of 1989 and the beginning of 1990, Eurogenetics N. V., Tessenderlo, Belgium, introduced to the European market two automated enzyme immunoassay analysers, the AIA-1200 and the AIA-600, produced, together with the reagents to run on them, by the Tosoh Corporation, Tokyo, Japan.

At the beginning of 1991, Eurogenetics started with a European Quality Control Programme for all users of these instruments. The aim was to keep the analytical variation of the analyses as low as possible. This article will review the results of the first 6 runs of this Quality Control Programme (over a time span of 12 months), and for reference purposes it will compare them with the results of the national Italian QC Programme, which uses the same type of samples at the same frequency and with the same statistical data reduction (10).

Methods

The AIA-instruments are based on the random accessibility concept, established through the standardization of procedures and the re-use of the calibration curves, and on the walk-away automation principle based on the minimization of manual intervention. As a tool to reduce analytical variation, the above mentioned Quality Control Programme was initiated. It is supervised by an independent international committee (with members from Italy, The Netherlands, France and Belgium).

The Quality Control Programme is an external quality survey incorporating AIA-users from all over Europe. This external quality control programme, meant as a link between the users, is organized in the same way as most national quality control programmes (8, 9). Lyophilized pooled serum samples are obtained from Technostandard, Brescia, Italy. The participants are supplied with a set of 9 of these sera (comprising three concentrations each of thyroid, fertility and tumour markers), mailed at room temperature. Over a period of 12 months, 6 pools of human serum are distributed, each pool being presented maximally 3 or 4 times, and in random combination with the others. The participants can therefore never know or guess the target concentrations of the samples. The same source of samples and the same distribution procedure was used in the Italian national QC Programme (10).

Participants performed the assays on their AIA-instrument and returned the results to their local Eurogenetics office or dealer, identifying the reagent lot numbers and the date of the last calibration, with a view to explaining any outliers that might occur.

The collected results were computer-processed, following the methods described by Pilo et al. (10). Briefly, the consensus mean, calculated from all the results after outlier rejection, is implicitly indicated to the participants as a target. The rejection of outliers is performed according to the method of Healy (11), and applied to the distribution of all results, using a symmetrical 5% trimming. Comparison of the deviation of the result produced by each laboratory (% deviation from the consensus mean) with the spread of all results (%CV) allows the evaluation of the acceptability of the analytical performance. A report containing this information is sent back to all participating laboratories.

Statistical significances in table 1 were calculated using the mean of the differences at $p < 0.5$ (12).

Results and Comments

The first run was started in May 1991. By now (January 1993), 8 runs have been performed. The same sample pool was presented repeatedly during a one year period (May 1991 to April 1992), and it was observed that the variation of the consensus means is smaller than the analytical variation of the respective value.

In table 1 we have summarized the first 6 runs of the external Quality Control Programme.

Thyroid markers

Thyrotropin: for sample 1, the %CV decreases significantly between run 2 and run 6, between run 4 and run 6 and between run 5 and run 6. For samples 2, 3 and 4 there are no significant changes of %CV. For sample 5 there is a significant decrease in %CV between run 1 and run 3, and between run 1 and run 5, but not between run 3 and run 5 (increase not significant). For sample 6 the %CV increases between run 2 and run 3.

Triiodothyronine: there are no statistical differences.

Thyroxine: there are no significant differences for samples 1 and 6. In sample 2, the %CV increases significantly between runs 1 and 3, but not between runs 1 and 5. For sample 3, a significant decrease occurs between runs 2 and 6. In sample 4, a significant increase exists between runs 1 and 4. For sample 5 the %CV decreases between run 1 and 5, and between run 3 and 5, but not between runs 1 and 3.

Fertility markers

Lutropin: there are no shifts for samples 2, 3 and 6. Sample 1 differs significantly in %CV between runs

Tab. 1. Overview of the first 6 results in the Eurogenetics Quality Control Programme (May 1991 – April 1992). For each run the number of participants (n) per quantity is given. For the 6 pools (Sample 1–6) over the 6 runs, the consensus mean, standard deviation (S.D.) and the % coefficient of Variation (C.V.) are given. In the last column the % half biological variation ($\frac{1}{2}$ BV) is given together with the literature reference (Ref.).

a. Thyroid markers

Thyro- tropin mIU/l	Run (n)		2 (56)		3 (56)		4 (70)		5 (84)		6 (94)		$\frac{1}{2}$ BV (%)
	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	
Sample no.	1 (56)												9–17 Ref. 7, 13, 15, 19, 22
1	1.10	0.18 16.6	0.26	0.07 25.4	1.09	0.12 11.2	0.27	0.08 28.5	0.26	0.08 30.6	0.28	0.06 21.2	
2			1.54	0.11 7.3			1.54	0.17 10.8	0.10	0.11 10.2	1.13	0.14 12.8	
3	3.29	0.32 9.9					3.15	0.29 9.2			1.53	0.15 9.9	
4	8.93	0.92 10.3			8.44	0.64 7.5			8.53	0.71 8.3			
5			17.47	1.32 7.6	16.70	1.48 8.9							
6													

Triodo- thyronine nmol/l	Run (n)		2 (58)		3 (58)		4 (72)		5 (80)		6 (86)		$\frac{1}{2}$ BV (%)
	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	Consen- sus Mean	S.D. %	
Sample no.	1 (58)												4–8 Ref. 2, 7, 13, 19
1	1.86	0.31 16.6	2.09	0.48 23.2	1.97	0.32 16.4	2.00	0.25 12.0	2.03	0.29 14.3	2.10	0.34 16.5	
2			2.60	0.48 18.4			2.63	0.30 10.9	1.92	0.28 14.6	1.91	0.28 14.5	
3	3.60	0.46 13.0			4.43	0.46 10.6	3.70	0.41 11.3			2.54	0.38 15.3	
4	4.30	0.54 12.7			5.65	0.54 9.6			4.42	0.51 11.4			
5			5.73	0.74 12.8									
6													

Tab. 1. ctd.

Thyroxine nmol/l	Run (n)		$\frac{1}{2}$ BV (%)												
	1 (56)		2 (57)		3 (56)		4 (66)		5 (70)		6 (70)		4-7 Ref. 7, 13, 18, 19, 20, 21		
Sample no.	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %
1				96.0	10.7	11.7	96.23	0.68	10.1	96.0	11.6	12.0	94.7	9.29	9.8
2	89.0	8.12	9.2				90.17	9.50	10.6				91.6	8.77	9.6
3				119.2	11.6	9.7				121.9	11.1	9.1			
4	200.6	17.9	8.9							214.8	29.4	13.7		10.3	9.0
5	241.2	23.0	9.5				262.8	25.9	9.9				240.5	23.20	9.2
6				315.4	42.5	13.8	312.9	17.4	5.6						

b. Fertility markers

Lutropin IU/l	Run (n)		$\frac{1}{2}$ BV (%)												
	1 (64)		2 (64)		3 (64)		4 (66)		5 (92)		6 (92)		7 Ref. 7		
Sample no.	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %
1				8.39	1.38	16.5	8.60	2.17	25.3	7.79	1.40	17.9			
2	31.02	3.43	11.0				31.32	3.94	12.6	30.84	3.65	11.8			
3				52.39	4.28	8.2	52.64	6.50	12.3	51.89	4.46	8.6			
4	27.85	3.75	13.5							27.62	2.83	10.2	30.48	5.40	17.7
5	18.40	3.20	17.4							18.64	2.09	11.2	20.88	3.30	15.8
6				8.54	1.42	16.6				8.56	1.39	16.2	9.52	2.07	21.7

Folli-
tropin

IU/l	Run (n)		$\frac{1}{2}$ BV (%)												
	1 (62)		2 (62)		3 (62)		4 (64)		5 (92)		6 (94)		15 Ref. 7		
Sample no.	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %	Consen- sus Mean	S.D.	C.V. %
1				27.0	2.68	9.9	25.86	3.39	13.1	26.40	3.46	13.1			
2	19.82	3.02	15.2				19.16	2.17	11.3	20.14	2.36	11.7			
3				74.81	8.38	11.2	70.77	8.62	12.2						
4	37.23	4.66	12.5							70.07	7.22	10.3	37.71	3.73	9.9
5	22.82	3.03	13.5							35.90	3.68	10.3	23.07	2.12	9.2
6				7.84	0.84	10.8				21.42	2.02	9.4	7.84	0.83	10.6

Eur. J. Clin. Chem. Clin. Biochem. / Vol. 31, 1993 / No. 12

Eur. J. Clin. Chem. Clin. Biochem. / Vol. 31, 1993 / No. 12

Tab. 1. ctd.

Ferritin ng/ml	Run (n)		2 (60)		3 (62)		4 (64)		5 (92)		6 (94)		$\frac{1}{2}$ BV (%)	
	Consen- sus	S.D.	Consen- sus	S.D.	Consen- sus	S.D.	Consen- sus	S.D.	Consen- sus	S.D.	Consen- sus	S.D.	18 Ref.	17
Sample no.	Mean	C. V. %	Mean	C. V. %	Mean	C. V. %	Mean	C. V. %	Mean	C. V. %	Mean	C. V. %		
1	32.94	3.74	30.33	3.98	31.11	3.03	31.40	3.02	33.86	4.27	34.07	3.94	11.6	
2	66.90	5.91			64.42	7.59			68.49	6.71	68.15	8.25	12.1	
3	81.59	7.63			116.70	7.72	81.18	7.01	124.30	11.9				
4		9.4				6.6								
5	117.70	12.2	10.3				217.80	21.10	9.6		230.80	23.70	10.3	
6	219.30	22.4	10.2											

Human β -chorionic gonado- tropin IU/l	Run (n)		2 (46)		3 (46)		4 (42)		5 (62)		6 (72)		$\frac{1}{2}$ BV (%)	
	Consen- sus	S.D.	Consen- sus	S.D.	Consen- sus	S.D.	Consen- sus	S.D.	Consen- sus	S.D.	Consen- sus	S.D.	— Ref.	—
Sample no.	Mean	C. V. %	Mean	C. V. %	Mean	C. V. %	Mean	C. V. %	Mean	C. V. %	Mean	C. V. %		
1	15.05	3.15	7.37	1.49	6.91	0.89	15.40	2.11	16.72	2.26	8.03	0.83	10.3	
2	23.68	4.60			22.19	2.46			25.28	3.46	24.02	5.19	21.6	
3	31.89	5.91			41.36	4.67	34.03	3.45	45.21	6.33				
4		18.5				11.3								
5	42.17	6.91	16.4				61.33	6.13	10.0		64.10	6.92	10.8	
6	59.57	9.90	16.6											

2 and 3 (increase) and between runs 3 and 5 (decrease). For sample 4, the %CV increases between run 1 and 6, and between runs 4 and 6. In sample 5, the %CV decreases between runs 1 and 6, but increases between runs 4 and 6.

Follitropin: for sample 1, the %CV increases between runs 2 and 3. The %CV for sample 2 decreases between runs 1 and 3, and between runs 1 and 5. For sample 3, there is an increase in %CV between runs 2 and 3, and between runs 2 and 4, but no significant difference between runs 3 and 4. Sample 4 shows no differences. The %CV for sample 5 decreases between runs 1 and 4, 1 and 6, but not between 4 and 6. The %CV for sample 6 increases between runs 2 and 5, but not between 2 and 6. Between runs 5 and 6, there is a significant decrease.

Prolactin: no differences can be observed.

Tumour markers

α -Foetoprotein: a very significant decrease in %CV can be seen in all combinations for all 6 samples from run 1 to run 6. Only sample 6 shows an increase in %CV between runs 4 and 6.

Carcinoembryonic antigen: sample 1 shows decrease in %CV between runs 2 and 6 and between runs 3 and 6, but not between 2 and 3. The %CV for sample 2 decreases from runs 1 to 4 and from 1 to 5. For samples 3, the consensus mean tends to decline, while the %CV decreases from run 1 to runs 3, 5 and 6. Between runs 3 and 5, the %CV increases, but between runs 5 and 6 it decreases, while no significant difference is seen between runs 2 and 6. Sample 4 shows no significant differences. Sample 5 shows a decrease in %CV between runs 2 and 3, but an increase between runs 2 and 5, and 3 and 5. Sample 6 shows a decrease in %CV between runs 2 and 4 and between runs 4 and 6.

Ferritin: no differences occur for samples 1 and 2. Sample 3 shows an increased %CV between runs 1 and 3, and between runs 3 and 6, but not between runs 1 and 6. Sample 5 shows a decrease in %CV from run 2 to run 3 and from run 2 to run 5. There is a concentration shift for sample 6, run 6, but no significant change in %CV.

Human β -chorionic gonadotropin: The %CV decreases in samples 1 and 2, but increases in sample 3 (from runs 3 to 5 and from runs 3 to 6). In sample 4, it decreases again from run 1 to run 4. The %CV for sample 5 decreases from run 2 to run 5, but increases

from runs 3 to 5. Sample 6 shows a decrease in %CV from runs 2 to 6. From run 4 to run 6 there is a concentration increase.

It can be concluded that the CV of some analytes was relatively high during run 1 and run 2, but decreased afterwards, as can be seen for e.g. thyrotropin, follitropin, α -foetoprotein, carcinoembryonic antigen and human β -chorionic gonadotropin. The decrease in CV is not so apparent for thyroxine, lutropin and ferritin, and even absent for triiodothyronine and prolactin. After runs 5 and 6 the CV's tend to stabilize.

After the 1st and the 2nd run, the international committee advised Eurogenetics to give certain instructions to the users. These were related to:

1. rinsing the wash probes more accurately;
2. replacing the wash probes after 2000 tests; and
3. rinsing the substrate container and tubings with ethanol.

Moreover, in a few cases, after consulting the users, a service intervention was made. In one case an instrument was even replaced.

In our view, the company, in taking this action, correctly carried out the advice of the committee, as can be seen in the evolution of the CV during the subsequent runs (tab. 1).

Also, towards the end of the observed period, larger batches of reagents were produced.

As stated in the introduction, the acceptability criteria for the analytical variation should be equal or lower than $\frac{1}{2}$ of the biological variation ($\frac{1}{2}$ BV). This quantity includes a measure of the intra-individual variation and the analytical variation (2, 7, 13–22). As can be seen in table 1, the analytical variations for thyrotropin, follitropin, prolactin, α -foetoprotein and ferritin are acceptable when compared with the $\frac{1}{2}$ BV. Carcinoembryonic antigen is almost in line, while triiodothyronine, thyroxine and lutropin still need some adjustment. For human β -chorionic gonadotropin, no data on $\frac{1}{2}$ BV were available in the literature.

Table 2 shows the comparison between the mean %CV of the Eurogenetics Quality Control Programme over the 6 first runs, and the mean %CV of the Italian National Quality Assurance Programme which uses the same samples and statistics as the Eurogenetics QC Programme. The CV given for the AIA and the other techniques is the mean as calculated from the European Quality Control Programme and from the Italian National QC Programme, respectively.

Tab. 2. Comparison between the mean %CV obtained by the AIA-1200 and AIA-600 users in the European Quality Control Programme and the mean %CV obtained by other techniques in the Italian national QC Programme, together with the % half biological variation ($\frac{1}{2}$ BV) References cited in tab. 1).

	Triiodo- thyronine	Thyroxine	Thyrotropin	Lutropin	Follitropin	Prolactin	Ferritin	Carcino- embryonic antigen	α -Foeto- protein	Human β -chorionic gonadotropin
	%CV	%CV	%CV	%CV	%CV	%CV	%CV	%CV	%CV	%CV
Delfia	—	—	—	21.7	15.1	17.1	—	—	19.5	—
Amerilite	8.5	—	9.2	13.9	18.5	14.0	8.0	11.7	13.5	—
IMX	8.7	7.3	11.3	8.5	7.4	15.5	11.8	10.0	15.3	—
ES300/ES600	11.3	8.3	10.5	18.7	9.5	11.3	11.2	—	—	—
AIA	14.1	10.1	13.7	14.7	11.5	14.4	10.3	10.1	11.1	14.7
DPC	13.3	11.3	21.9	45.2	25.4	20.2	20.5	—	—	23.8
BYK	10.1	9.2	15.9	—	—	—	—	—	—	—
Amerlex	12.4	6.8	17.8	—	—	—	11.7	—	—	—
Serono	13.9	8.4	13.3	22.6	11.4	14.1	—	12.9	13.7	16.5
CIS	—	13.7	16.2	14.7	18.4	15.3	—	18.0	12.9	—
Sorin	—	—	—	—	—	—	—	—	20.9	—
% $\frac{1}{2}$ BV	4–8	4–7	9–17	7	15	20	18	9	11	—

From table 2, it can be concluded:

Thyroid markers

The reagents for triiodothyronine and thyroxine from almost all companies give variations that are equal or higher than the $\frac{1}{2}$ BV. The IMX method is clearly superior to the others for these quantities. For thyrotropin, one sees that the non-isotopic methods show a much lower CV than the isotopic ones.

Fertility markers

For lutropin, the IMX and AIA give the lowest CV values, but there is still room for improvement with respect to the $\frac{1}{2}$ BV. For follitropin, the AIA-method gives an average quality. Again IMX is the best one, while Boehringer and AIA are $\leq \frac{1}{2}$ BV. Generally, all methods have a CV above the $\frac{1}{2}$ BV for lutropin and approximately equal to the $\frac{1}{2}$ BV for follitropin. AIA, IMX and Serono give almost equivalent values for prolactin, which are $< \frac{1}{2}$ BV. There is no difference in the results of isotopic and non-isotopic methods, except for DPC.

Tumour markers

The AIA gives the lowest CV for carcinoembryonic antigen and α -fetoprotein, and second lowest for ferritin. Human β -chorionic gonadotropin again gives an average CV. In terms of $\frac{1}{2}$ BV, the AIA-method is the best for carcinoembryonic antigen and α -fetoprotein. No company meets the limits for carcinoembryonic antigen, but AIA is very close.

In general, except for triiodothyronine and thyroxine, it can be seen that the non-isotopic methods tend to give lower CV values than the isotopic methods.

As stated in the introduction, there are two main manufacturer-dependent sources of analytical variation: lot-to-lot variation and test performance, which is related in the case of the AIA to the instrument and the reagent quality. Both of these aspects were taken into consideration by enlarging the batch size, by an accurate interpretation of the results and by advising specific actions to the customers after the first two runs of the Eurogenetics Quality Control Programme. As a result, the total %CV decreased from run 3 onwards. Further adaptations are needed, however, for some analytes to meet $\frac{1}{2}$ BV specifications.

Eurogenetics confirmed to us that short term reagent improvements (for triiodothyronine, thyroxine, lutropin and carcinoembryonic antigen) are imminent.

Moreover, the Quality Control Programme has been extended to include more analytes as from 1993, i.e. insulin, prostate-specific antigen, prostatic acid phosphatase, IgE, somatotropin and free thyroxine. The observation time for these quantities has been too short for sound conclusions. The first impression, however, is that the tendency is the same for these recent quantities as for the other more extensively investigated quantities.

We consider it important to participate in a national Quality Control Programme in order to see how the methodology one uses is related to that of other manufacturers. Also, participation in an international Quality Control Programme with the same methodology provides an important overview of the intrinsic

quality of the method used, as long as internationally accepted reference methods are not available.

We conclude that the organization of a Quality Control Programme by Eurogenetics for its own apparatus and reagents is a daring, but highly advisable operation. In the future, every company that commercializes such instruments, should organize an international QC Programme.

Acknowledgement

We thank Dr. G. C. Zucchelli, CNR, Pisa, Italy for the data from the Italian national Quality Control Programme and Dr. F. E. L. Lecock, Eurogenetics N.V., Belgium, for skilful assistance during the preparation of this paper.

References

- Harris, E. K., Kanofsky, P., Shakarji, G. & Cotlove, E. (1970) Biological and analytical components of variations in long term studies of serum constituents in normal subjects. II. Estimation biological components of variation. *Clin. Chem.* 16, 1022–1027.
- Costongs, G. M. P. J., Janson, P. C. W., Bas, B. M., Hermans, J., Van Wersch, J. W. J. & Brombacher, P. J. (1985) Short term and long term individual variations and critical differences of clinical chemical laboratory parameters. *J. Clin. Chem. Clin. Biochem.* 23, 7–16.
- Fraser, C. G. (1988) The application of theoretical goals based on biological variation data in proficiency testing. *Arch. Pathol. Lab. Med.* 112, 404–415.
- Williams, G. Z., Young, D. S., Stein, M. R. & Cotlove, E. (1970) Biological and analytical components of variations in long term studies of serum constituents in normal subjects. I. Objectives, subject selections, laboratory procedures and estimation of analytical deviation. *Clin. Chem.* 16, 1016–1021.
- Cotlove, E., Harris, E. K. & Williams, G. Z. (1970) Biological and analytical components of variations in long term studies of serum constituents in normal subjects. III. Physiological and medical implications. *Clin. Chem.* 16, 1028–1032.
- Young, D. S., Harris, E. K. & Cotlove, E. (1970) Biological and analytical components of variations in long term studies of serum constituents in normal subjects. IV. Results of a study designed to eliminate long term analytical deviations. *Clin. Chem.* 17, 403–410.
- Ricos, C. & Arbos, M. A. (1990) Quality goals for hormone testing. *Ann. Clin. Biochem.* 27, 353–358.
- Wellcome Immunoassay Quality Control Programme. Wellcome Research Laboratories, Beckenham, U.K.
- Sufi, S. B. & Goncharov, N. (1984) Standardization of radioimmunoassays, for use in WHO Collaborative Studies. In: *Cost Benefit and Predictive Value of Radioimmunoassay* (Albertini, A., Ekins, R. P. & Galen, R. S., eds.) pp. 39–48, Elsevier, Amsterdam.
- Pilo, A., Zucchelli, G. C., Chiesa, M. R., Masimi, S. & Clerico, A. (1991) The CNR external quality assessment program for immunoassays: Statistical analysis and reports for participants. *Ann. Inst. Super. Sanita* 27, 469–478.
- Healy, M. J. R. (1979) Outliers in clinical chemistry quality control schemes. *Clin. Chem.* 25, 675–677.
- Swinsow, T. D. V. (1982) *Statistics at Square One*. British Medical Association, pp. 1–60, London.
- Hölzel, W. G. E. & Deschmer, W. (1988) Intra-individual variation of serum thyroxine and triiodothyronine in pregnancy. *Clin. Chem.* 34, 2063–2065.
- Winkel, P., Bentzon, M. W., Statland, B. E., Mouridsen, H. & Sheike, O. (1982) Predicting recurrence in patients with breast cancer from cumulative laboratory results: A new technique for the application of time series analysis. *Clin. Chem.* 28, 2057–2067.
- Ponchaux, C. (1989) Variations spontanées des concentrations sériques des hormones thyroïdiennes et thyroïdienne au cours du nyctémère. Thèse pour le doctorat en médecine, Lille.
- Reiner, C. B., Smith, S. J. & Wells, T. W. (1982) The U.S. national reference preparation for alpha-fetoprotein in mid-pregnancy maternal serum. *Clin. Chem.* 28, 709–716.
- Sizaret, P., Tuyns, A., Martel, N., Jouvenceaux, A., Levin, A., Ong, Y. W. & Rive, J. (1975) Alpha-fetoprotein levels in normal males from seven ethnic groups with different hepatocellular carcinoma risks. *Ann. N. Y. Acad. Sci.* 257, 136–155.
- Skendrel, L. P., Barnett, R. N. & Platt, R. (1985) Medically useful criteria for analytical performance of laboratory tests. *Am. Journ. Clin. Pathol.* 82, 200–205.
- Browning, M. C., Ford, R. P., Callaghan, S. J. & Fraser, C. G. (1986) Intra- and inter-individual biological variation of five analytes used in assessing thyroid function: Implications for necessary standards of performance and the interpretation of results. *Clin. Chem.* 32, 962–966.
- Linnet, K. (1989) Choosing quality control systems to detect maximum clinically allowable analytical errors. *Clin. Chem.* 35, 284–288.
- Williams, G. Z., Widowson, G. M. & Penton, J. (1978) Individual character of variation in time series studies of healthy people. II. Differences in values for clinical chemical analytes in serum among demographic groups by age and sex. *Clin. Chem.* 24, 313–320.
- Feldt-Rasmussen, U., Hyloft Petersen, P., Blaabjerg, O. & Hørder, M. (1980) Long-term variability in serum thyroglobulin and thyroid related hormones in healthy subjects. *Acta Endocrinol.* 95, 328–334.

Guido M. P. J. Costongs and Piet C. W. Janson
Maaslandziekenhuis
Afdeling Klinische Chemie
Walramstraat 23
NL-6131 BK Sittard
The Netherlands

